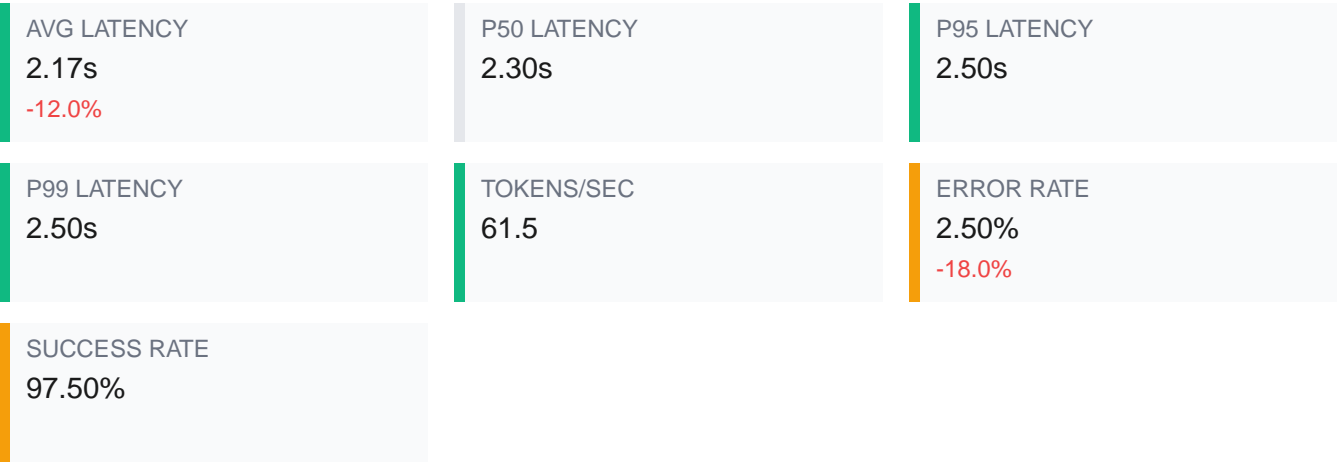


Engineering Performance Report

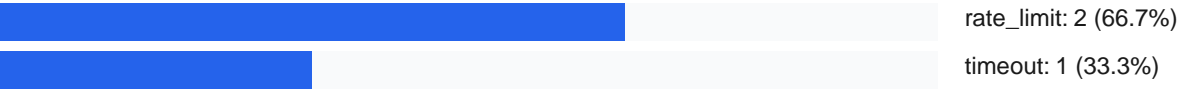
November 30, 2025 - December 14, 2025

Performance Metrics



Error Analysis

Error Distribution (3 total)



Tool Performance

Tool Name	Calls	Success Rate	Avg Time	Failures
web_search	3	66.7%	1943ms	1
dataset_manager	2	50.0%	2100ms	1
analytics_export	1	100.0%	2500ms	0
training_control	1	100.0%	800ms	0
token_analyzer	1	100.0%	600ms	0

5 unique tools analyzed

Model Comparison

Model	Messages	Total Tokens	Total Cost	Cost/Msg	Tokens/Msg
gpt-4	3	42,000	\$2.10	\$0.7000	14000
claude-3-sonnet	1	11,000	\$0.44	\$0.4400	11000

2 models used

Technical Recommendations

[MEDIUM] Improve Tool Reliability

Tool failure rate at 25.0%

Impact: Reliable tools reduce fallbacks and improve response quality

- Add timeout handling to tool implementations
- Implement retry logic with exponential backoff
- Add better error messages for debugging
- Consider circuit breaker patterns for external APIs

[MEDIUM] Reduce Token Usage

Cost per message (\$0.0880) is above threshold

Impact: Token optimization can reduce costs by 30-50%

- Audit prompts for unnecessary verbosity
- Implement prompt compression techniques
- Use smaller models for simpler tasks
- Cache common responses where appropriate